AN ABSTRACT OF THE THESIS OF

Yibing Jia for the degree of Master of Science in Botany and Plant Pathology presented on April 26, 1996. Title: Transcript Analysis of *FELDMANNIA Sp.* Virus, FsV: Characterization of the Major Capsid Protein Gene and Its Relationship to Known Viruses

Redacted for Privacy

Abstract approved:

Russel H. Meints

The *Feldmannia sp.* virus is a large icosahedral virus that persistently infects marine brown alga *Feldmannia sp.*. So far, there is no information available about viral genome replication, gene structure and gene expression in this unique viral-host system. The purpose of this study was to characterize the general features of viral transcripts in the virus producing sporophyte plants. Northern analysis, using four cosmid clones that cover the entire viral genome, showed that there were six major transcripts and at least eighteen minor transcripts in the virus producing sporophyte plants. These transcripts are not evenly distributed in the viral genome. A 5.7 kb *Bam*HI fragment - R was found to encode a 1.5 kb and a 0.9 kb major transcript, and those two major transcripts were chosen for detailed sequence analysis. The 1.5 kb transcript was identified as the putative major capsid protein (MCP) gene. The FsV MCP has significant similarity with the major capsid protein of *Chlorella* virus-PBCV-1 and with iridoviruses, fish lymphocystis disease virus, frog virus 3, and with African swine fever virus.

Transcript Analysis of *Feldmannia Sp.* Virus, Fsv: Characterization of the Major Capsid Protein Gene and Its Relationship to Known Viruses

by

Yibing Jia

A THESIS

submitted to

Oregon State University

in partial fulfillment of
the requirement of the degree of

MASTER OF SCIENCE

Completed April 26, 1996
Commencement June, 1996

Master of Science thesis of Yibing Jia presented on April 26, 1996

APPROVED:

Redacted for Privacy

Major professor, representing Botany and Plant Pathology

Redacted for Privacy

Chair of Department of Botany and Plant Pathology

Redacted for Privacy

Dean of Graduate School

I understand that my thesis will become part of the permanent collection of Oregon State University libraries. My signature below authorizes release of my thesis to any reader

Redacted for Privacy

Yibing Jia, Author

Typed by: Yibing Jia

ACKNOWLEDGMENTS

I would like to thank the following people for their help and encouragement:

Dr. Russel Meints for allowing me the opportunity to pursue my education, and his advice during my graduate studying.

Dr. Stella Coakley, chair of the Department of Botany and Plant Plant Pathology, for her kind advice.

Dr. Eric Henry and Mr. Richard Ivey, who gave me help when I was working on my experiments.

Dr. Reg McParland, Barbara Robbins, Anne-Marie Girard of the OSU, Center for Gene Research and Biotechnology, Central Service Laboratory, for their aid in obtaining DNA sequence data and for oligonucleotide synthesis.

And finally my wife Cassie (Chunzhi) Dong for giving me the love.

TABLE OF CONTENTS

LIST OF FIGURES

FIGURE Page

Transcript Analysis of *Feldmannia Sp.* Virus, FsV: Characterization of the Major Capsid Protein Gene and Its Relationship to Known Viruses

1. INTRODUCTION

Brown algae (Phaeophyceae) are widely distributed intertidal inhabitants of sea coasts around the world. Since the early 1970s, viruses or viruslike particles (VLPs) have been reported in several brown algae (Phaeophyceae) on the basis of transmission electron microscopy. The viruses or VLPs have been found in various life history stages of brown algae. These include swimming zoospores (Baker and Evans 1973), newly settled zoospores (Toth and Wilce 1972, Oliveria and Bisalputra 1978), vegetative cells (La Claire and West 1977, Oliveria and Bisalputra 1978), and developing sporangia and gametangia (Clitheroe and Evans 1974, Markey 1974, Müller *et al.,* 1990). Henry and Meints (1992) reported a persistent virus that occurs intracellularly in a filamentous marine brown algae of the genus *Feldmannia.* This or a related virus has similarly been reported in *Feldmannia simplex* and *Feldmannia irregularis* (Müller and Frenzer 1993, Friess-Klebl *et al.,* 1994). Additionally, viruses have been shown in other Ectocarpales, i.e., *Ectocarpus fasciculatus* (Clitheroe and Evans 1974), *Ectocarpus siliculosus* (Müller *et al.,* 1990), *Streblonema uvaeformis* (La Claire and West 1977) and *Sorocarpus sp.* (Oliveira and Bisalputra 1978). As interest in brown algae (Phaeophyceae) as economically important species harvested for their constituent colloids (e.g. alginate) develops, pathogens that present a problem in high density culture will need to be studied. Very little is known about the significance of

viruses as pathogens in field algal populations, or about their individual host ranges (reviewed by Van Etten *et al.*, 1991, Henry and Meints 1992, Müller 1992; Müller and Stache 1992, Müller and Parodi 1993).

1.1 General characteristics of *Feldmannia sp.* and FsV

Plants of the genus *Feldmannia* are uniseriate, branched, filamentous forms that develop single-celled unilocular (meiotic) sporangia to produce haploid zoospores. The unilocular sporangia of the field isolate reported by Henry and Meints (1992, ostensibly a diploid sporophyte) did not develop normally. They were initiated normally, as lateral protrusions from cells of the branched filaments. The protrusions swelled, became ovoid, and were then separated from the mother cell by formation of a crosswall. The nucleus in the protrusion divided, but the multinucleate cytoplasm did not cleave to form zoospores, and the sporangium lost pigmentation rather than becoming more darkly pigmented. Spores were not observed in the sporangia but rather numerous polyhedral particles formed in their place. Transmission electron microscopy revealed the particles to be large icosahedral viruses called *Feldmannia sp.* virus -- FsV (approximately 1-5 x 10^6 per sporangial cell). No trace of these particles was evident in vegetative cells of sporophyte and gametophyte plants, nor has infectious entry of viruses into sporangia been observed (Henry and Meints 1992). When placed under unialgal or axenic culture conditions abundant formation of virus-

producing sporangia can be obtained by regulation of culture medium, light and culture density.

Structurally, FsV is a large polyhedral particle approximately 150 nm (calculated) in diameter, with a electron-dense, spherical core that is surrounded by a monolayer shell. One-dimension SDS-PAGE analysis of FsV virion protein showed only about eight structural proteins ranging in size from 14 kDa to 110 kDa (unpublished data). The polypeptide with an apparent molecular weight of about 49 kDa was the most abundant structural protein (unpublished data). Because it is difficult to obtain large quantities of virion particles, other components (like lipid and sugar) in the virion are not known. Preliminary analysis of the genome structure of the FsV has been reported (Henry and Meints 1992). Two genome size-classes (158 and 179 kb, with multiple variants of each) were found, whose individual abundance in viral preparations is affected by culture temperature (Ivey *et al.,* in press). Viral DNAs isolated from algal cultures grown at low temperature (10 °C) is predominantly of the large size class. At intermediate temperatures (15 °C) both genomes are present in roughly equal proportions while at high temperature (18-20 °C) replication of the small genome size class predominates. The viral genome was mapped as a double-stranded circular DNA genome (Ivey *et al.,* in press). Both the large and small viral genome contain a 173 bp direct repeats in large numbers (Lee *et al.,* 1995). Additional studies of a *Hin*dIII-*Sst* I 4.5 kbp subclone demonstrated the presence of three transcriptionally active open reading frames, one of which contains an ATP binding site and a "RING" zinc finger motif (Krueger *et al.,* 1996).

1.2 The purpose of this research

The molecular fine structure, mechanisms of replication, infectivity, and gene expression in brown algal viruses have only recently begun to be studied. But so far, no transcript analysis of brown algal viruses has been reported. In this research I examined FsV transcription from the virus genome of *Feldmannia* sporophytes. Four cosmid clones (C1-08, C1-24, C1-30 and C1-49) which cover the entire viral genome were used as probes in Northern hybridization analysis to characterize FsV viral transcripts. After characterization of the viral genome transcripts, I chose two major transcripts which were located in the viral genome *Bam*HI fragment - R to do further sequence analysis. One of the major transcripts encodes the major capsid protein. Sequence analysis of the gene demonstrated a significant homology to other viruses in the databanks including *Chlorella* viruses PBCV-1 (Graves and Meints 1992, Van Etten *et al.,* 1995), fish lymphocystis disease virus (Schnitzler, Darai,. 1993), iridescent viruses (Stohwasser *et al.,* 1993, Cameron 1990, Tajbkhsh *et al.,* 1990), and the African Swine Fever virus (Lopez-Otin *et al.,* 1990).

2. FURTHER LITERATURE REVIEW

A virus infected *Feldmannia simplex* (Phaeophyceae) was isolated and characterized by Friess-Klebl *et al.,* in 1994. The virus infected plants grew normally but formed elongated, weakly pigmented vesicles instead of normal zoidangia. The virus particles are about 120-150 nm in diameter with an electron-dense core that was surrounded by a monolayered shell. The virus particles were only seen in zoidangia and not in vegetative cells. The virus genome was characterized as circular dsDNA with a size of 210-240 kb. The virion contains one major polypeptide of MW 55 kDa and at least six additional polypeptides (MW=15 - 120 kDa). These features are similar to the FsV in our lab, except the viral genome size is bigger than FsV's. So far there is no further report on the *Feldmannia simplex* virus.

Some progress also has been made in the understanding of viruses infecting *Ectocarpus siliculosus,* an alga closely related to *Feldmannia* organism in which a virus (EsV) has been demonstrated. It contains a very large (350 kbp) dsDNA genome also circular in topology and displaying Mendelian inheritance (Müller *et al.*, 1990; Müller *et al.*, 1991; Lanka *et al.*, 1993, Müller *et al.*, 1993). Klein *et al.*,(1995) characterized the largest of the three glycoproteins of *Ectocarpus siliculosus* virus -- gp-1 which has an apparent molecular weight of 60 kDa. The gp-1 protein, suggested by the authors to be a coat protein, is a 661 amino acid polypeptide and has a calculated molecular weight of 72 kDa. The discrepancy between the calculated and apparent molecular weight was explained as the results of proteolytic processing in the maturation of the gp-1 protein.

Also the gp-1 protein is rich in hydrophilic amino acids. Although viral infection virus-free *Ectocarpus siliculosus* plants system has been established, there has been no report on viral gene transcription and expression.

3. MATERIALS AND METHODS

3.1 Culture condition

The virus-infected alga was collected at Oaro, New Zealand, in March 1984. The alga was brought into culture by Dr. Eric Henry and is presently maintained in the enriched seawater medium 2 x PES (McLachlan 1973) at 15-20 °C under continuous light with vigorous aeration.

3.2 RNA isolation

Algal filaments were collected by filtration through a nylon screen (95 μm), 10 g of fresh weight algae were ground with a mortar and pestle with 5g sterilized sea sand and liquid nitrogen. The ground powder was transferred to a 250 ml centrifuge tube that contained 50 ml of CTAB extraction buffer (Chang *et al.*, 1993) (5 ml/g fresh weight), extracted two times with an equal volume of chloroform:isoamyl alcohol (24:1), and the phases were separated at 10,000 rpm at room temperature in a Sorvall GSA rotor. One-fourth volume of 10M LiCl was added to the supernatant and mixed. The RNA was precipitated at 4° C overnight and harvested by centrifugation at 10,000 rpm for 20 min in a Sorvall SA600 rotor. The supernatant was discarded and the pellet was dissolved in 3 ml of SSTE buffer, extracted once with an equal volume of chloroform:IAA, and 2 volumes of ethanol were added to the supernatant, followed by

precipitation of the RNA at -20°C for at least 2 hours and a spin of 20 min. in a SA600 rotor to pellet the RNA. The supernatant was removed and the pellet vacuum-dried. The pellet was resuspended in autoclaved DEPC-treated deionized H_2O. The concentration of RNA samples were quantified by absorbency reading at 260 nm and was visualized on a denatured agarose gel. The RNA samples were denatured at 72 °C for 10 min. and quenched in ice. The samples were loaded onto a 1.2% agarose gel containing 6% formaldehyde and 1x MOPS buffer. Electrophoresis was performed at 2 volts/cm for 22 hours.

3.3 Northern blotting analysis

For Northern blot analysis, equal amounts of total RNA (4 μg/lane) were loaded onto a MOPS denaturing agarose gel for electrophoresis. The electrophoresed RNAs were transferred overnight from the denaturing gel onto a nylon Hybond-N membrane (Amersham, Arlington Heights, IL) with 20x SSC buffer. The membrane was baked for 2 hours at 80°C in a vacuum-oven. The blots were pre-hybridized with formamide hyb-solution for 1 hour at 42 °C, then hybridized with ^{32}P labeled probes overnight at 42 °C. After hybridization, the membrane was washed twice with 0.1x SSC and 1% SDS for 20 min at 42 °C. The hybridized blots were wrapped with plastic film and autoradiographed with an intensifying screen at -80 °C. All Northern blot hybridization experiments were repeated twice to confirm the results.

3.4 Radioactive labeling

Double-stranded DNA probes were radiolabeled with (α-^{32}P) dATP by a random primed extension reaction (Multiprimer DNA labeling systems, Amersham, Arlington Heights, IL). Oligonucleotide probes were radiolabeled with (γ-^{32}P) ATP by T4 polynucleotide kinase (New England Biolabs, Beverly, MA). Unincorporated nucleotides were removed by Sephadex G-50 or G-25 spin column chromatography .

3.5 Cloning and sequencing

The transcripts containing *Bam*HI-fragment R were digested with *Eco*RI and ligated into pUC118/119 using T4 DNA ligase (New England Biolabs, Beverly, MA) and transformed into the *E. coli* host JM101 for sequencing.

The template DNAs (double-stranded plasmid DNA) for sequencing were purified with a QIAprep (Qiagen Inc., Chatsworth, CA) plasmid DNA purification kit. Sequencing was performed with an M13 forward primer and custom-synthesized oligonucleotide primers on an ABI 373A DNA sequencer. DNA sequences were compiled and analyzed using the GCG program set (Devereux *et al.*, 1984). Sequence similarity searches with GenBank and EMBL databases were performed using programs based on the BLAST algorithm (Altschul *et al.*, 1990).

3.6 Transcript mapping

The 5′ end of the transcription start sites were mapped by T4 DNA polymerase on cloned viral genomic DNA and total *Feldmannia sp.* RNA (Hu and Davidson 1986); 0.5μg of ssDNA, 20 μg of *Feldmannia sp.* total RNA and $1x10^4$ -10^5 cpm of ^{32}P-labeled primer were used.

3.7 3′RACE (rapid amplification cDNA 3′ end)

3′ RACE experiment was used to amplify the 3′ end of viral mRNA. One micro-gram of *Feldmannia sp.* total RNA and 20 pmol of oligo-dT was used to perform first strand synthesis with M-MLV-RT reverse transcriptase (Gibco BRL, Gaithersburg, MD). After the reverse transcription reaction, a gene specific primer (20 pmol) was used to perform PCR amplification.

3.8 Isolation of polyadenylated RNA

The polyadenylated RNA was isolated with an Oligotex-dT mRNA Mini Kit (Qiagen, Chatsworth, CA) as described by the manufacture. Five hundred to seven hundred micro-gram of total RNA were used to isolate polyadenylated RNA.

4. RESULTS

4.1 Northern characterization of FsV genome from algal sporophytes

Synchronous infection of the algal host -- *Feldmannia sp.* by FsV is not yet possible hence study of enriched viral-induced messages is not possible. To initiate transcriptional studies, total RNA isolated from the virus-infected sporophyte cultures of *Feldmannia sp.* at the time of maximal virus production was used. It is likely that only the most abundance messages will be detected by this approach. Total RNA was electrophoresed on denaturing agarose gels and transferred to Hybond-N nylon membrane. The Northern blots were probed with cosmids containing viral DNA that represented the entire viral genome. Four cosmid clones (C1-08, C1-24, C1-30,C1-49) sufficient to cover the entire viral genome, were used as probes as shown in Figure 2. The size of the transcripts hybridizing to each of the cosmid clones was estimated (shown in Table). The transcripts sizes ranged from 0.9 kb to 5.5 kb. Six major transcripts and at least eighteen minor transcripts were identified by the Northern analysis (Fig. 2). Northern analysis also indicated that viral transcriptional activity is not evenly distributed throughout the viral genome. After further Northern screening with *Bam*HI fragments of C1-24, a 5.7 kb *Bam*HI fragment R was found that contains the two major transcripts of 1.5 kb and 0.9 kb. Figure 3 shows the Northern blot of total RNA isolated from *Feldmannia sp.* sporophyte plants probed with fragment R.

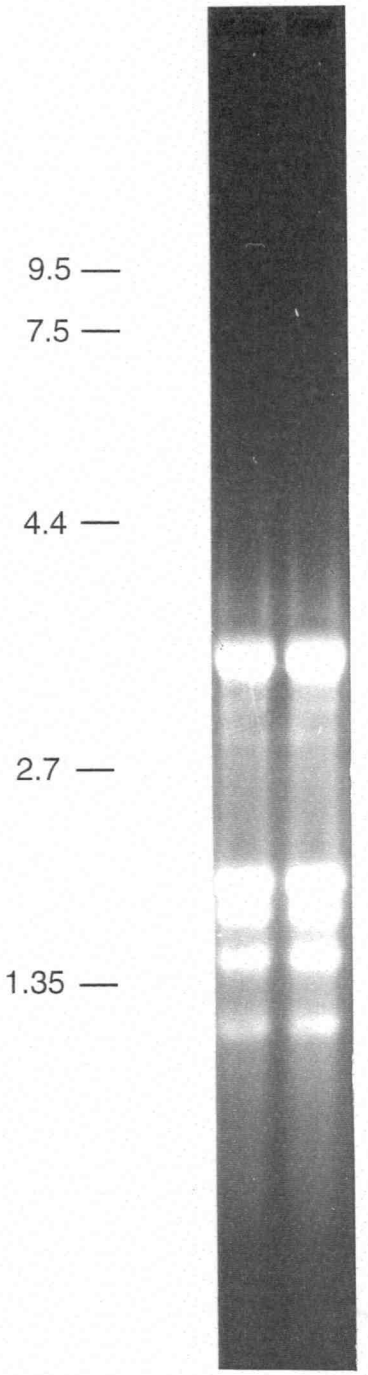

Fig. 1 Total RNA isolated from virus infected *Feldmannia sp.* sporophyte

Table

Size of the transcripts of FsV

	Cosmid			
	C1-08	C1-24	C1-30	C1-49
Size of transcripts (kb)	5.0	5.5	3.4	3.4
	3.4	4.4	2.7	2.7
	2.7	3.4	2.5	2.1
	1.6	3.2	1.6	1.6
	1.1	3.0	1.5	1.5
		2.5	1.3	
		2.4	1.2	
		2.1	1.0	
		1.5		
		0.9		

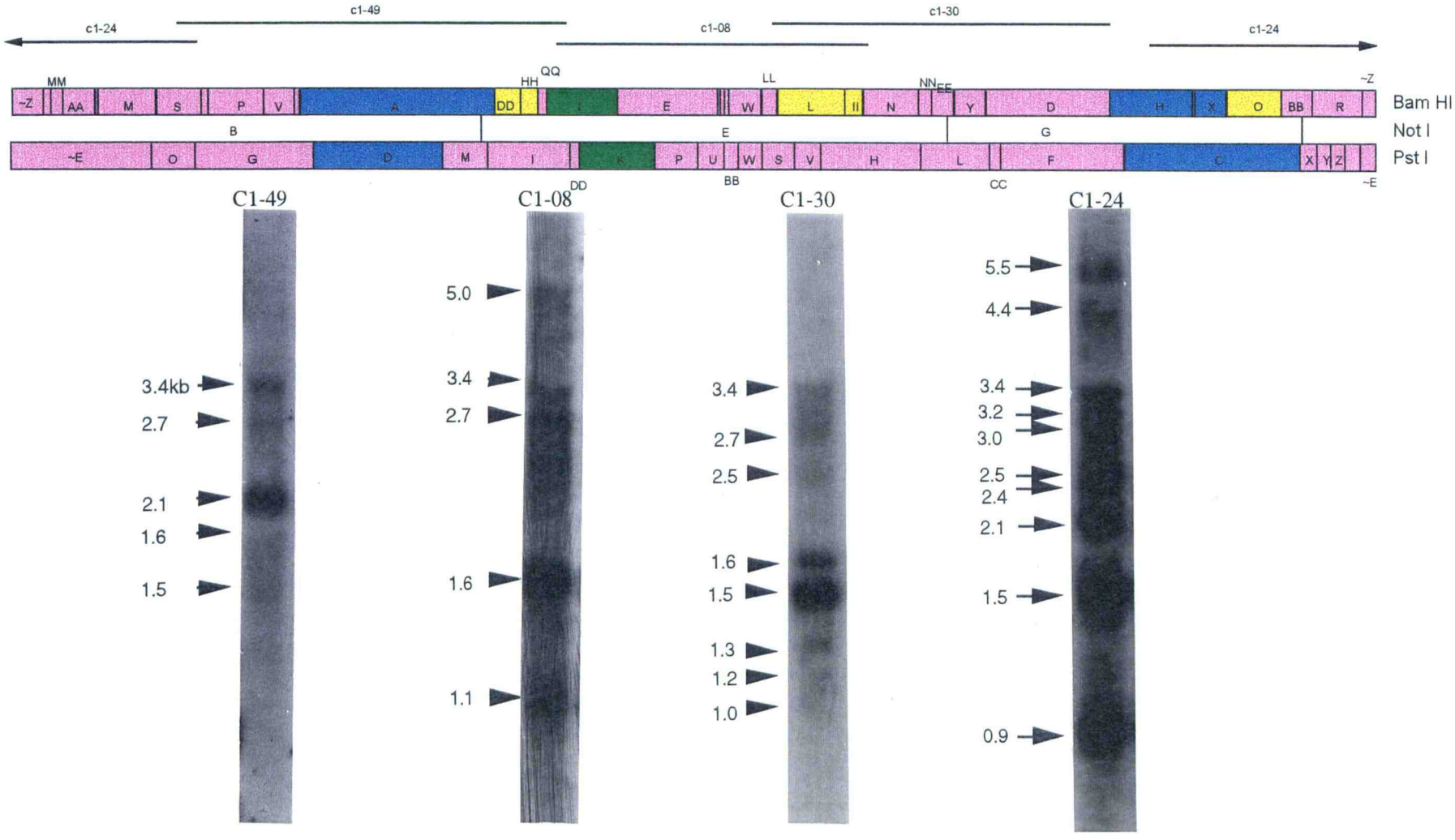

Fig. 2 Map of the FsV viral genome. Northern blots probed with four cosmid clones

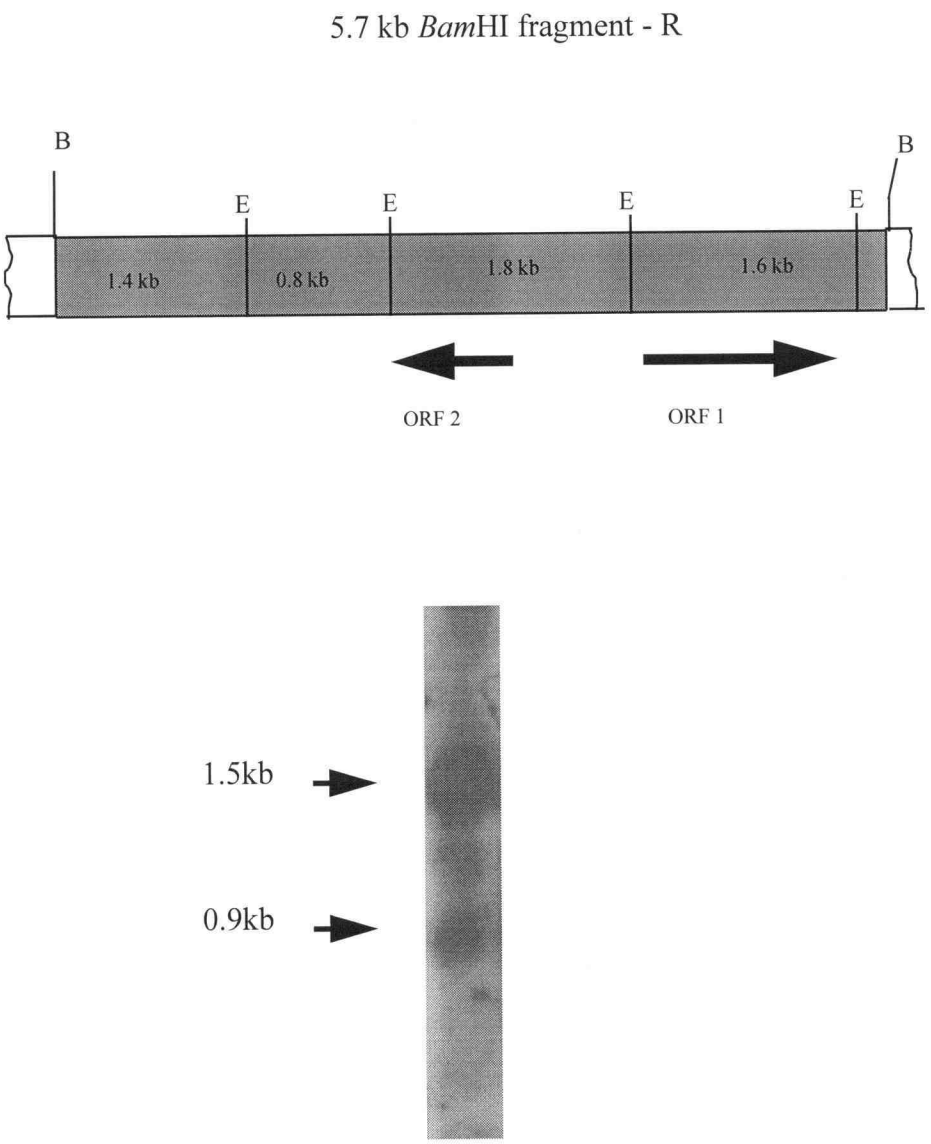

Fig. 3 Map of the 5.7 kb *Bam*HI fragment - R and the Northern blot probed with R, the ORF 1 and ORF 2 are shown as heavy arrows.

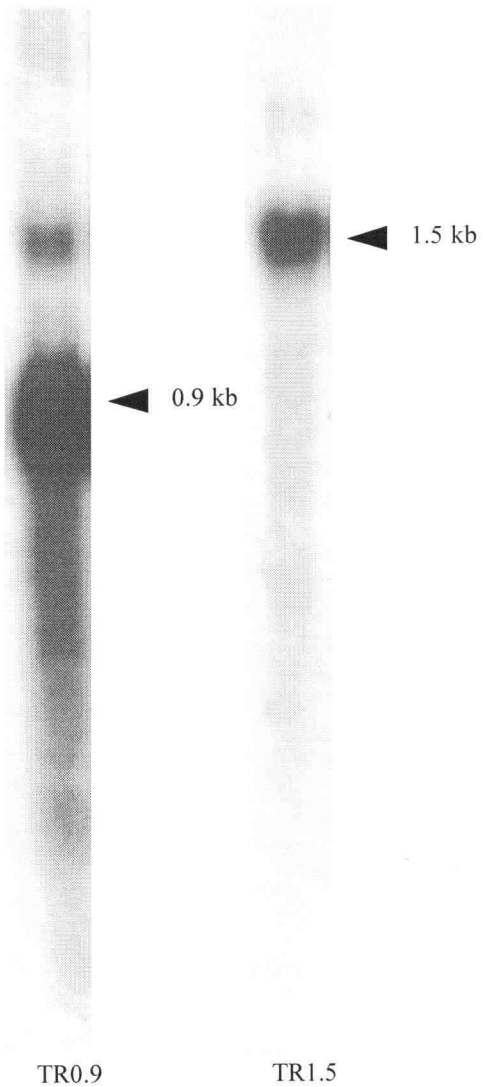

Fig.4 Northern blots probed with oligonucleotides: TR1.5 and TR0.9 probes

4.2 Sequence analysis

The 5.7 kb *Bam*HI fragment R, which contains the 1.5 kb and 0.9 kb major viral transcripts, was digested by *Eco*RI and subcloned into plasmid pUC118/119 for DNA sequencing. DNA sequence analysis identified two open reading frames (ORFs) of 1308 bp (ORF 1) and 597 bp (ORF 2). The two ORFs are in opposite orientations. The restriction map of the R fragment and the orientation of the ORFs of the two major transcripts are shown in Fig. 3. Generally, there is no obvious A+T rich region upstream of the AUG codon. The common eukaryotic polyadenylation signal AAUAAA was found in ORF 1, but the results of 3′ RACE experiment and Northern blotting analysis of polyadenylated RNA of *F. sp.* showed that the mRNA of ORF 1 was not polyadenylated. Northern blotting analysis of polyadenylated RNA of *F. sp.* also showed that the mRNA of ORF 2 was not polyadenylated. There is no classic TATA box upstream of the first initiation codon AUG in either ORF. The sequence TTTTTNT, which has been shown to be important for transcription termination of vaccinia virus early genes and Chlorella virus genes, was not found in either ORF.

To confirm that these two ORFs are the 1.5 kb and 0.9 kb transcripts that have been identified by Northern blotting analysis, two oligonucleotide probes, TR1.5 and TR0.9, which are complimentary to the sense strand and close to the 5′-end first AUG codon (for TR1.5 from +15 to +59, for TR0.9 from +12 to +56) were used to probe a Northern blot of *Feldmannia sp.* total RNA. The results (shown in Figure 4.) were

consistent with the Northern blots probed with the entire R fragment. Therefore the 1308 bp ORF 1 represents the 1.5 kb viral transcript, and the 597 bp ORF 2 represents the 0.9 kb transcript.

The peptide sequences deduced from the two ORFs were compared to GeneBank databases and analyzed by the FASTA and BLAST algorithms, and sequence similarities were compared. Fig. 5 shows the amino acid sequence deduced from ORF 1; Fig. 6 shows the amino acid sequence deduced from ORF2 . ORF 1 has three possible asparagine-linked glycosylation sites (Asp-X-Ser/Thr) which have been found in the PBCV-1 major capsid protein, Vp54, that has 10 % lipid component (Skrdla *et al.*, 1984), but ORF 2 does not contain such sites. The calculated molecular weight of the ORF 1 product is 48.8 kDa. SDS-PAGE analysis of FsV virion protein showed a major protein band of molecular weight of 49 kDa. Consequently we concluded that the ORF1 encodes the major capsid protein gene of *Feldmannia sp.* virus. The database search found no significant match to ORF 2. The function of this gene is unknown.

The 5´end mapping assay showed that the putative major capsid protein gene has multiple-transcription start sites. The two mapped transcription start sites — AACT sequence are located 138 nt and 158 nt, respectively, 5′ to the first AUG. The 138 nt sequence 5′ to the first AUG is the major transcription start site. Around the transcription initiation sites there are no AT rich regions, nor TATA box.

```
        10                  30                  50
AAAGTCGTCAAGCTCGGAGGCTCGGTTCCGATCGACGGCGAGTCGCCTGTGCTCATCCCG
        70                  90                 110
TTCTTGGAGACGGGGGGCAGCGCCCAGACGGCGTCATTGACGCTGTCGGACAACTCTTCC
       130                 150                 170
GTGACCTTGACGTACGATGAGACCGCCGGCACCATCGGGATCGGTGGGGTCGACTACTCT
       190                 210                 230
CCCGGCGAATCCACGATTGTCGATTCCCACAAGCTGACCCTTGTCGAGGTTTAGATATTT
       250                 270                 290
AGATATTTAGAAATTTTGAAATGTTGTAATTCGAATTTGAATTTGAATTTGAATTCGACT
       310                 330                 350
TCGTCAAGACAGATCGATCTCGACCAACGGGTGTCCCGACGATTCGTCGTCAAATATACT
       370                 390                 410
GATGGACGGTTCGCTGTATTCCTGGCGAATCTCCCGATGGGGACGCCACCAACAACAACA
       430                 450                 470
ATAGTAGCTCAAACACATGAGTAATACACTACTCGTGTATTACTCATGGCACAGTTTTTA
       490                 510                 530
GTGCTCGTTGTGGGTACATTTCTGTAATCCCCCTGAAGTAGTTGGGATTGACATACTGGC
       550                 570                 590
GAGCCTGCTCTTGTTCTCGCATGTCAATGTACCTCCTCATCCTCGCGTCGTGGGCATTCA
       610                 630                 650
TTCGATCCACGTTTCGCTGGTCGAAGGCGTTGGTCTGGAAGGCCCTGTTGTAGGACGAGT
       670                 690                 710
ACTCGTTCTGGAAACGCTGGTCTTGGCGCTTAAGGTTTTCCTCGTACCTCCTGACGAAAT
       730                 750                 770
TCCCGGCGTCGTCGTACTTTTTCGCCAGGGCAAGGGCCTTTCCGCGAGGCGTTTTGATCG
       790                 810                 830
AGATGCACTTGAGAGTCGCCGGATCTAAAATCTTTCCCGCCCCACAGGACTTGACGCGAC
       850                 870                 890
AGTTCCCGAACTTGTCGCGTTCCAACCACGGCAGCGGCCTGCCCCCGTTGTCGAAGCAGG
       910                 930                 950
GCATGCCCTCGCCACAGTGCCCGTCGTAGTCCATCTCCTTGTCGATCGAGCATCCTCTCT
       970                 990                1010
TCTTGATCACTTTCTTTTTCAGGACCTTCTTTGTGGGCTTTTTAGTTGTTTTCGGAGCCA
      1030                1050                1070
TTGGTTGAGTGCGTGTGAGCGTATGTCGACCTATCAACATAAAAATCTCAGTTTTTGACG
      1090                1110                1130
CTTTGACTACTTAACGGTGCGCGAGTCCGTACACACACGCGAAAAAAATCAGGAGGAGGA
      1150                1170                1190
GGGCAAGTTGGAATGTCATTCTGTTCCAACTGTCTCGGACACACCCCGTCTTCGTACACG
      1210                1230                1250
AGGCTCTATAGGATTCATATTTTTGGGAAACTTCGTCAAAGCCGGACTTTTTGTATACCA
      1270                1290                1310
CACCCAGCTTTTCGTTGACCATGTTGTGAATCTTCCACAGCCATCGAGTGAGTTGGTCGC
```

Fig.5 The sequence for part of 5.7 kb *Bam*HI fragment -R. The putative major capsid protein translate is presented. The asparagine glycosylation sites are bold and underlined

```
      1330                 1350                 1370
GGTTGGCGGCCTCTACGGGCAGTCTGCCGATGAACTCGTTGTAGGATGCTCTGCAAAGCC
      1390                 1410                 1430
TGCAGGGTAATACACTCCCGACCGTTTCGAAAAATATAGTATAATCGTTGACCTCTTCCA
      1450                 1470                 1490
CGGACGGCGTTGTTGGGAAACCGTGAGCAACGCTGTGCAAAAACTTCCACCCGGCTGGAC
      1510                 1530                 1550
CCCATTCTGACGTCGCGAGCCCGTTCATGGCGAGCGAAGAGTACTTTGGCATCTTTACGC
      1570                 1590                 1610
AAGATACAAATTACAAATATGAAACACTTTGTGCAGTTCCGGTGTTTTTAAATGTTGATG
      1630                 1650                 1670
GTGGACATACAAAATCATACAACATGCCTGCAGGTGGAGGGGTAAAGATTTGAATTGCCC
      1690                 1710                 1730
CTAAGAGGACGGTGCGGTGTCGCGAGGCTCCACTAACCACCCCGCAACGTGCGGGGGTAC
      1750                 1770                 1790
CGTTCTAGTCCGGGCGCTCTCGATCCTCATTGCATCTAGTAGGTGGGAGCGTCTAGTTCC
      1810                 1830                 1850
GGGTCCCTCCTGGAATGCGGCTCTCTCTCTCTCCCGAGAGCCAGGGCAAGGTTTGTGAAA
      1870                 1890                 1910
ACGGTCAAAGCTAATTCTCTTTTTAGGAGATTTGGTAAGACCGTCGGTGGGCCCGTGTCG
      1930                 1950                 1970
AAAGGCGCGGTGAACCATCGCTATCGACTGGGGCACGAATCGGTTCCTAAAGAGTCTTAG
      1990                 2010                 2030
AAACTCCATGGGGCTCAATGTACAGTCAACTCCCGGGTGAATCCGGGTAACTCCGGGCAA
      2050                 2070                 2090
AGCCCGGGCCTCGTGTGATCGGGCGGACTGAAAATCGAAACCCGGGAAGCACGAGGGTCT
      2110                 2130                 2150
CGTTGCAAAACGTGAATTCAAATCGCAGCCGTTGGTCGCCAAAATGTACACCTCAACGGG
      2170                 2190                 2210
AACCCGGACATGACCCTCTTCAAGACCGTACACAAGAGGTACACGTCGTTCGCGGAAGAC
 <MCP>   M  T  L  F  K  T  V  H  K  R  Y  T  S  F  A  E  D
      2230                 2250                 2270
CTGGAGGAGAACGACTTCAGCGCGGGAACGGTCGGCTTCGGGCAGAAAGTGTCGGCGAAC
L  E  E  N  D  F  S  A  G  T  V  G  F  G  Q  K  V  S  A  N
      2290                 2310                 2330
GTCTCCAGGTATGGGGACCTCGTGACCGACATGTTCATGGAGGTCGCCCTCCCGCCGATC
V  S  R  Y  G  D  L  V  T  D  M  F  M  E  V  A  L  P  P  I
      2350                 2370                 2390
GAGGCCGCGGCCACCGTGACCAACGCCGAAGGGGCCGAGGTCGCCGACGCCGACAAGGCC
E  A  A  A  T  V  T  N  A  E  G  A  E  V  A  D  A  D  K  A
      2410                 2430                 2450
GCGTACTGGGTCAACGCCATCGGGTACGCTCTTATCTCGGAGATTCAGATCGAGATCGGG
A  Y  W  V  N  A  I  G  Y  A  L  I  S  E  I  Q  I  E  I  G
      2470                 2490                 2510
GGAACCGAGGTAGACACGCTCTACCCCGAGTGGATGTTCTTCTGGGAAAGTATGACCCAG
G  T  E  V  D  T  L  Y  P  E  W  M  F  F  W  E  S  M  T  Q
      2530                 2550                 2570
AGACCCGGGGCGCGGCTCGGCGAGCAGATCGGGAAGTTCGCGTACTCCGCCGACGTGGAA
R  P  G  A  R  L  G  E  Q  I  G  K  F  A  Y  S  A  D  V  E
```

Fig.5 (continued)

```
      2590                2610                2630
GAGGACATGATCGAGTTCGCCCAGCAGGCCCGTACGCTGTACGTCCCGCTCCCGTTCTGG
E  D  M  I  E  F  A  Q  Q  A  R  T  L  Y  V  P  L  P  F  W
      2650                2670                2690
TTCAACAAGTACTTCATGGAGACCGGGCTCAGCATTCCCCTCATCGCCCTGACCTACCAC
F  N  K  Y  F  M  E  T  G  L  S  I  P  L  I  A  L  T  Y  H
      2710                2730                2750
GAGATCAAGGTCAAGGTGACGTTCCGCCCGCTGTCGGAGTGCTGTTGCGTGGTTTACCGC
E  I  K  V  K  V  T  F  R  P  L  S  E  C  C  C  V  V  Y  R
      2770                2790                2810
GCGGAGGACGAGACCCACGGGGAGTACTTCGCCCTCGCCGAGGGGAAGACCCCCGTGAAC
A  E  D  E  T  H  G  E  Y  F  A  L  A  E  G  K  T  P  V  N
      2830                2850                2870
ACCACGTCCGGGAGCACCCTCGTGTCCTCCGACATGGACGCGAAGCTTCTCATCTCGTAC
T  T  S  G  S  T  L  V  S  S  D  M  D  A  K  L  L  I  S  Y
      2890                2910                2930
GTCTACCTCGACAAGGCCGAGCGAGACGCCTTCGCGTCGACCGAGCACTCGTACCTGATC
V  Y  L  D  K  A  E  R  D  A  F  A  S  T  E  H  S  Y  L  I
      2950                2970                2990
ACGAGGACGCAGAGGCAGCTGCACGCCATCACTTCGGCCGGGTCGGCCTCGGACCAGATC
T  R  T  Q  R  Q  L  H  A  I  T  S  A  G  S  A  S  D  Q  I
      3010                3030                3050
AAGCTGTACTTCAACCACCCGTCGAACTGCCTCGCGTGGTTCGTGCGACCCACGGACTGG
K  L  Y  F  N  H  P  S  N  C  L  A  W  F  V  R  P  T  D  W
      3070                3090                3110
ACGACAAACCGTCGTCGTTTCTCCGTGGGGCACATGGACTCGTTCGACTTCTCCCTGCAC
T  T  N  R  R  R  F  S  V  G  H  M  D  S  F  D  F  S  L  H
      3130                3150                3170
ACCGACTCGGACGTGTCCGTTTGGGGAGACGTGATCGACCCCGTCAAGTCAGCGTCCCTG
T  D  S  D  V  S  V  W  G  D  V  I  D  P  V  K  S  A  S  L
      3190                3210                3230
AACCTGAACGGTCACAGCAGGTTCCCGGACGGCATGCCAGGGCTATTCTTCCGACAGACT
N  L  N  G  H  S  R  F  P  D  G  M  P  G  L  F  F  R  Q  T
      3250                3270                3290
CAGCCCATCATGAAGTGGCCGAACTGTTCGGACGGATTCATGTACGTGTTCTCGTTCTCC
Q  P  I  M  K  W  P  N  C  S  D  G  F  M  Y  V  F  S  F  S
      3310                3330                3350
CTGCAGGGAGGCGCGTGGCAACCGACATCGACCCTCAACATGTCGCGCATCGACCACGTG
L  Q  G  G  A  W  Q  P  T  S  T  L  N  M  S  R  I  D  H  V
      3370                3390                3410
CAGCTCGAGCTCAAGTACGGGTCGAACATCCCCACGTCCGACGTGTTTGTTTTCGCAGAG
Q  L  E  L  K  Y  G  S  N  I  P  T  S  D  V  F  V  F  A  E
      3430                3450                3470
TCGTACAACCTCCTCGTGGTGAAGGATGGCATGGGGGGCGTGAGGTACAGTAATTAAGTA
S  Y  N  L  L  V  V  K  D  G  M  G  G  V  R  Y  S  N  *
```

Fig.5 (continued)

```
      3490                 3510                 3530
GCGGAACGAGTGATAAAACAAGAAATGATATCGTTTTTCACTCCCAATCCCAATAAAATT
      3550                 3570                 3590
TTAATAACAATATAAATTATTTCTCCACCTTGAAGTCCCAGTCGCCCAGACGAAGGAGAC
      3610                 3630                 3650
ACTTGGACACCTTACACCGCGAGTAAAACACGGACGTGTCGATCGAGTCGTGCAGCAAGA
      3670                 3690                 3710
CTCGCTTGAGCCAGATGTTGATCGAACACTTGTTTTCAAAAGCGACGTGGGTCACACAGC
      3730                 3750                 3770
CGTAGATGGCTCGCAGATCGTGGAAACGAGTAGATACCGAATTCGTAATAATCTTGGCGG
      3790                 3810                 3830
CGTCGCCGAGTTTGTGCGAAAACCGAGCGTTGATGACGGCGTTCGAGTCGTCGGGCCCGA

TCGGGATCC
```

Fig.5 (continued)

```
AACTTGCCCTCCTCCTCCTGATTTTTTTCGCGTGTGTGTACGGACTCGCGCACCGTTAAG
TAGTCAAAGCGTCAAAAACTGAGATTTTTATGTTGATAGGTCGACATACGCTCACACGCA
                                M  L  I  G  R  H  T  L  T  R  T
CTCAACCAATGGCTCCGAAAACAACTAAAAAGCCCACAAAGAAGGTCCTGAAAAAGAAAG
  Q  P  M  A  P  K  T  T  K  K  P  T  K  K  V  L  K  K  K  V
TGATCAAGAAGAGAGGATGCTCGATCGACAAGGAGATGGACTACGACGGGCACTGTGGCG
  I  K  K  R  G  C  S  I  D  K  E  M  D  Y  D  G  H  C  G  E
AGGGCATGCCCTGCTTCGACAACGGGGGCAGGCCGCTGCCGTGGTTGGAACGCGACAAGT
  G  M  P  C  F  D  N  G  G  R  P  L  P  W  L  E  R  D  K  F
TCGGGAACTGTCGCGTCAAGTCCTGTGGGGCGGGAAAGATTTTAGATCCGGCGACTCTCA
  G  N  C  R  V  K  S  C  G  A  G  K  I  L  D  P  A  T  L  K
AGTGCATCTCGATCAAAACGCCTCGCGGAAAGGCCCTTGCCCTGGCGAAAAAGTACGACG
  C  I  S  I  K  T  P  R  G  K  A  L  A  L  A  K  K  Y  D  D
ACGCCGGGAATTTCGTCAGGAGGTACGAGGAAAACCTTAAGCGCCAAGACCAGCGTTTCC
  A  G  N  F  V  R  R  Y  E  E  N  L  K  R  Q  D  Q  R  F  Q
AGAACGAGTACTCGTCCTACAACAGGGCCTTCCAGACCAACGCCTTCGACCAGCGAAACG
  N  E  Y  S  S  Y  N  R  A  F  Q  T  N  A  F  D  Q  R  N  V
TGGATCGAATGAATGCCCACGACGCGAGGATGAGGAGGTACATTGACATGCGAGAACAAG
  D  R  M  N  A  H  D  A  R  M  R  R  Y  I  D  M  R  E  Q  E
AGCAGGCTCGCCAGTATGTCAATCCCAACTACTTCAGGGGGATTACAGAAATGTACCCAC
  Q  A  R  Q  Y  V  N  P  N  Y  F  R  G  I  T  E  M  Y  P  Q
AACGAGCACTAAAAACTGTGCCATGAGTAATACACGAGTAGTGTATTACTCATGTGTTTG
  R  A  L  K  T  V  P  *
AGCTACTATTGTTGTTGTTGGTGGCGTCCCCATCGGGAGATTCGCCAGGAATACAGCGAA
CCGTCCATCAGTATATTTGACGACGAATCGTCGGGACACCCGTTGGTCGAGATCGATCTG
```

Fig.6 The sequence of the 0.9 kb transcript (ORF2).

4.3 The major capsid protein gene and its relationship to other viruses

The deduced amino acid sequence of the FsV putative major capsid protein was compared with the sequences of the major capsid proteins of *Chlorella* virus-PBCV, *Chilo* iridescent virus, *Tipula* iridescent virus, iridescent virus 22, frog virus 3, fish lymphocystis disease virus, African swine fever virus and Ectocarpus virus. Figure 7 shows the multiple alignment results. Twenty-seven percent amino acid identity found between FsV and *Chlorella* virus PBCV-1. The major capsid protein of FsV possesses similarity with iridoviruses in several conserved regions. All of these major capsid proteins have molecular weight of approximately 50 kDa, except the major capsid protein of ASFV which has a molecular weight of 72 kDa.

A dendrogram (Fig. 8) showing the relatedness among those of major capsid proteins, which was generated with GCG program that aligns multiple sequences and compares the relatedness among FsV and the other viruses. The dendrogram grouped all iridoviridae viruses into one clade, algal viruses into one clade that is more close related to iridoviridae and a separated clade of ASFV.

The database search also indicated that the major capsid protein of FsV contains an in-frame thymidine kinase like sequence. Flügel *et al.*, (1982, 1985) reported that a thymidine kinase activity has been shown to be associated with the purified virion of fish lymphocystis disease virus (Iridoviridae), but the function of the thymidine kinase activity in the viral life cycle is unknown.

```
           1                                                                                                    70
   {fsv}   .......... .......... .......... ...MTLFKTV HKRYTSFAED LEENDFSAGT VGFGQKVSAN
  {irv22}  ..MSMSSSNI TSGFIDIATF DEIEKYMYGG PTATAYFVRE IRKSTWFTQV PVPLSRNTGN AAFGQEWSVS
  {tipuv}  ..MSMSSSNI TSGFIDIATF DEIEKYMYGG PTATAYFVRE IRKSTWFTQV PVPLSRNTGN AAFGQEWSVS
  {chilv}  ..MSISSSNV TSGFIDIATK DEIEKYMYGG KTSTAYFVRE TRKATWFTQV PVSLTRANGS ANFGSEWSAS
  {fislv}  .MTSVAGSSV TSAFIDLATY DTIEKHLYGG DSAVAYFVRE TKKCTWFSKL PVLLTRCSGS PNFDQEFSVN
  {frogv}  .MSSVTGSGI TSGFIDLATY DNLERAMYGG SDATTYFVKE HYPVGWFTKL PSLAAKMSGN PAFGQQFSVG
   {ecto}  .......... .......... .......... ......FKTV HRRYTNFAED FEENDFSAGT VGFGQRVSAN
   {pbcv}  .......... MAGGLSQLVA YGAQDVYLTG NPQITFFKTV YRRYTNFAIE SIQQTIN.GS VGFGNKVSTQ
   {asfv}(25)LNSRISNI KNVNKSYGKP DPEPTLSQIE ETHMVHFNAH FKPYVPIGFE YNKVRPHTGT PTLGNKLTFG
Consensus  ---------- ---------- ---------- ------F--- -R-YT-F--- --------G- --FGQ--S--

           71                                                                                                   140
   {fsv}   VSRYGDLVTD MFMEVALPP. ......IEAA ATVTNAEGAE VADADKAA.. .YWVNAIGYA LISEIQIEIG
  {irv22}  ISRAGDYLLQ TWLRVNIPP. ......VTL. .SGLLGNTYS LR........ ..WTKNLMHN LIREATITFN
  {tipuv}  ISRAGDYLLQ TWLRVNIPP. ......VTL. .SGLLGNTYS LR........ ..WTKNLMHN LIREATITFN
  {chilv}  ISRAGDYLLY TWLRVRIPS. ......VTLL STNQFGANGR IR........ ..WCRNFMHN LIRECSITFN
  {fislv}  VSRGGDYVLN SWMTVRIPA. ......IKLK ADNRMNNNGT IR........ ..WCKNLFHN LIKQTSVQFN
  {frogv}  VPRSGDYILN AWLVLKTPE. ......VELL AANQLGDNGT IR........ ..WTKNPMHN IVESVTLSFN
   {ecto}  VSRYGDLITD .......... .......... .......... .......... .......... ..........
   {pbcv}  ISRNGDLITD IVVEFVLTK. ......GGNG GTTYYPAE.. .......... .........E LLQDVELEIG
   {asfv}  IPQYGDFFHD MVGHHILGAC HSSWQDAPIQ GSSQMGAHGQ LQTFPRNG(36)YYCEYPGER LYENVRFDVN
Consensus  ISR-GDYIL- -WL-V--P-- ------V-L- ---------- -R-------- --W--N--HN LI-E--I-FN

           141                                                                                                  210
   {fsv}   GTEVDTLYPE WMFFWESMTQ RPGARLGEQI GKFAYSAD.. ..VEEDMIEF AQ.....QAR TLYVPLPFWF
  {irv22}  DLVAARFDNY HLDFWSAFTV PASKRNGYDN MIGNVSSL.. ..I..NPVAP GGTLGSVGGI NLNLPLPFFF
  {tipuv}  DLVAARFDNY HLDFWSAFTV PASKRNGYDN MIGNVSSL.. ..I..NPVAP GGTLGSVGGI NLNLPLPFFF
  {chilv}  DLVAARFDHY HLDFWAAFTT PASKAVGYDN MIGNVSAL.. ..IQPQPVPV APATVSLPEA DLNLPLPFFF
  {fislv}  DLVAQKFESY FLDYWAAFSM CGSKRAGYNN MIGNTIDM.. ..IQPVDHTG .....MLPEK VLVLPLPYFF
  {frogv}  DISAQSFNTA YLDAWSEYTM PEAKRTGYYN MIGNTSDL.. ..INPAPATG QDGARVLPAK NLVLPLPFFF
   {ecto}  .......... .......... .......... .......... .......... .......... ..........
   {pbcv}  GQRIDKHYND WFRTYDALFR MNDDR..... ..YNYRRM.. ..TDWVNNEL VG.....AQK RFYVPLIFFF
   {asfv}  GNSLDEYSSD VTTLVRKFCI PGDKMTGYKH LVGQEVSV.(75)VHYSCNGP QTPKYYQPPL ALWIKLRFWF
Consensus  D--------- ---FW-AFT- ---KR-GY-N -I---S---- --I------- ---------- -L-LPLPFFF

           211                                                                                                  280
   {fsv}   NKYFMETGLS IPLIALTYHE IKVKVTFRPL SECCCVVYRA EDETHGEYFA LAEGKTPVNT TS...GSTLV
  {irv22}  SR...DTGVA LPTAALPYNE MQINFNFRDW HELLILTNSA LVPPASSYVS IV........ VG...THISA
  {tipuv}  SR...DTGVA LPTAALPYNE MQINFNFRDW HELLILTNSA LVPPASPYVP IV........ VG...THISA
  {chilv}  SR...DSGVA LPTAALPYNE MRINFQFHDW QRLLILDNIA AVASQTVVPV VG........ AT...SDIAT
  {fislv}  SR...DSGVA LPSAALPYNE IRLTFHLRDY TELLIFQHKQ ....DCTIIP IT........ AA...DLEYG
  {frogv}  SR...DSGLA LPVVSLPYNE IRITVKLRAI HDLLILQHNT ....TGAISP IV........ AS...DLAGG
   {ecto}  .......... .......... .......... .......... .......... .......... ..........
   {pbcv}  NQ...TPGLA LPLIALQYHE VKL....... .......... .......YFT LASQVQGVNY NG...SSAIA
   {asfv}  NE...NVNLA IPSVSIPFGE RFITIKL.AS QKDLVNEFPG LFIRQSRFIP GRPSRRNIRF KPWFIPGVIS
Consensus  -------G-A LP--ALPYNE --I---FR-- -ELLIL---- -------Y-- I--------- ----------
```

Fig.7 Alignment comparison of the amino acid sequences of the major capsid proteins from iridoviruses, African swine fever virus, *Chlorella* virus (PBCV-1), *Ectocarpus* virus and *Feldmannia sp.* virus (bold letter).

```
          281                                                                 350
   {fsv}  SSDMDA.KLL ISYVYLDKAE RDAFASTEHS YLI....TR. ..TQRQLHAI TSAGSASDQI KLYFNHPSNC
 {irv22}  APVLGPVQVW ANYAIVSNEE RRRMGCAIRD ILI....EQV QTAPRQNYVP ..LTNASPTF DIRFSHAIKA
 {tipuv}  APVLGPVQVW ANYAIVSNEE RRRMGCAIRD ILI....EQV QTAPRQNYVP ..LTNASPTF DIRFSHAIKA
 {chilv}  APVLHHGTVW GNYAIVSNEE RRRMGCSVRD ILV....EQV QTAPRHVWNP ..TTNDAPNY DIRFSHAIKA
 {fislv}  KPDLKDVQVW ITNAVVTNEE RRLMGTTPRD ILV....EQV QTAPKHVFQP ..LTIPSPNF DIRFSHAIKL
 {frogv}  LPDTVEANVY MTVALITGDE RQAMSSTVRD MVV....EQV QAAPVHMVNP ..RNATTFHT DMRFSHAVKA
  {ecto}  .......... .......... .......... .......... .......... .......... ..........
  {pbcv}  GAAQPTMSVW VDYIFLDTQE RTRFAQLPHE YLI....EQL QFTGSETATP SATTQASQNI RLNFNHPTKY
  {asfv}  EISLTNNELY INNLFVTPEI HNLFVKRVRF SLIRVHKTQV THTNNNHHDE KLMSALKWPI EYMFIGLKPT
Consensus ---L----VW --Y--V---E R--+------ -LI----EQV Q-AP-----P ------S--- ---F-H--K-

          351                                                                 420
   {fsv}  LAW..... FVRPTDWTTN RRRFSVGHMD SFDFSLHTDS DVS..... .....VWGDV IDPVKSASLN
 {irv22}  LFFAVRNKTS AAEWSNYAT. .......... SSPVVTGATV NYE..... .....P.TGS FDPIANTTLI
 {tipuv}  LFFAVRNKTS AAEWSNYAT. .......... SSPVVTGATV NYE..... .....P.TGS FDPIANTTLI
 {chilv}  LFFAVRNTTF SNQPSNYTT. .......... AYPVLTSTTV ILE..... .....PSTGA FDPIHHTTLI
 {fislv}  LFFGVRNTTH AAVQSNYTT. .......... ASPVILEEAY ASD..... .....LSLVA ADPIANVTLV
 {frogv}  LMFMVQNVTH PSVGSNYTC. .......... VTPVVGVGNT VLE..... .....PAL.A VDPVKSASLV
  {ecto}  .......... .......... .......... .......... .......... .......... ..........
  {pbcv}  LAWNFNNPTN YGQYTALANI PGACSGAGTA AATVTTPDYG NTG..... .....TYNEQ LAVLDSAKIQ
  {asfv}  WNISDQNPHQ HRDWHKFGHV VNAIMQPSHH AEVSFQDRDT ALPDACSSIS DISPITYPIT LPIIKNISVT
Consensus L-F---N-T- ----S---T- ---------- ---V------ ---------- ---------- FDPI----LI

          421                                                                 490
   {fsv}  LNGHSRFPDG MPGLFFRQTQ PIMKWP...N CSDGFMYVFS FSLQG.GAWQ PTSTLNMSRI DHVQLELKY.
 {irv22}  YENTNRLGAM GSDYFSLINP FYHAPT...I PSFIGYHLYS YSLHF.YDLD PMGSTNYGKL TNVFVVPA..
 {tipuv}  YENTNRLGAM GSDYFSLINP FYHAPT...I PSFIGYHLYS YSLHF.YDLD PMGSTNYGKL TNVSVVPQ..
 {chilv}  YENTNRLNHM GSDYFSLVNP WYHAPT...I PGLTGFHEYS YSLAF.NEID PMGSTNYGKL TNISIVPT..
 {fislv}  YENSARLNEM GSEYYSLVQP YYFGGS...I PIETGYHMYC YSLNM.MDMD PMGSTNYGRL SNVSMKLK..
 {frogv}  YENTTRLPDM GVEYYSLVEP WYYATS...I PVSTGHHLYS YALSL.QDPH PSGSTNYGRL TNASLNVT..
  {ecto}  .......... .......... .......... .......... .......... .......... ..........
  {pbcv}  LNGQDRFA.T RKGSYFNKVQ PYQSIG...G VTPAGVYLYS FALKP.AGRQ PSGTCNFSRI DNATLSLTYK
  {asfv}  AHGINLIDKF PSKFCSSYIP FHIGGNSIKT PSSPGAMMIT FALKPREEYQ PSGHINVSRA REFYISWDTD
Consensus Y----R---- ----F----- ---------- ----G--L-S -SL------- P-G--N--RL -NV-------

          491                                                                 553
   {fsv}  .......... .......... .GSNIPTSDV FVFAESYNLL VVKDGMGGVR YSN*...... ...
 {irv22}  ...ASSAA.I SAAGGTGGQA GSDYAQSYEF VIVAVNNNIV RIENSLVRNR RRWSREGPMV MVC
 {tipuv}  ...ASPAA.I AAAGGTGGQA GSDYPQNYEF VILAVNNNIV RISGGETPQN ........YI AVC
 {chilv}  ...ASPAAKV GAAGTGPAGS GQNFPQTFEF IVTALNNNII RISGGALGFP VL........ ...
 {fislv}  ...TSDKAVV NAGGGGGNMS GYKDAQKFEF LTMAINHNVI RIKNGSMGFP VL........ ...
 {frogv}  ...LSAEATT AAAGGGGNNS GYTTAQKYAL IVLAINHNII RIMNGSMGFP IL*....... ...
  {ecto}  .......... .......... .......... .......... .......... .......... ...
  {pbcv}  TCSIDATSPA AVLGNTETVT ANTATLLTAL NIYAKNYNVL RIMSGMGGLA YAN....... ...
  {asfv}  YVGSITTA.. .......... ........DL VVSASAINFL LLQNGSAVLR YST*...... ...
Consensus ---------- ---G------ G--------- VV-A-N-NII -I--G--G-- ---------- ---
```

Fig.7 (Continued)

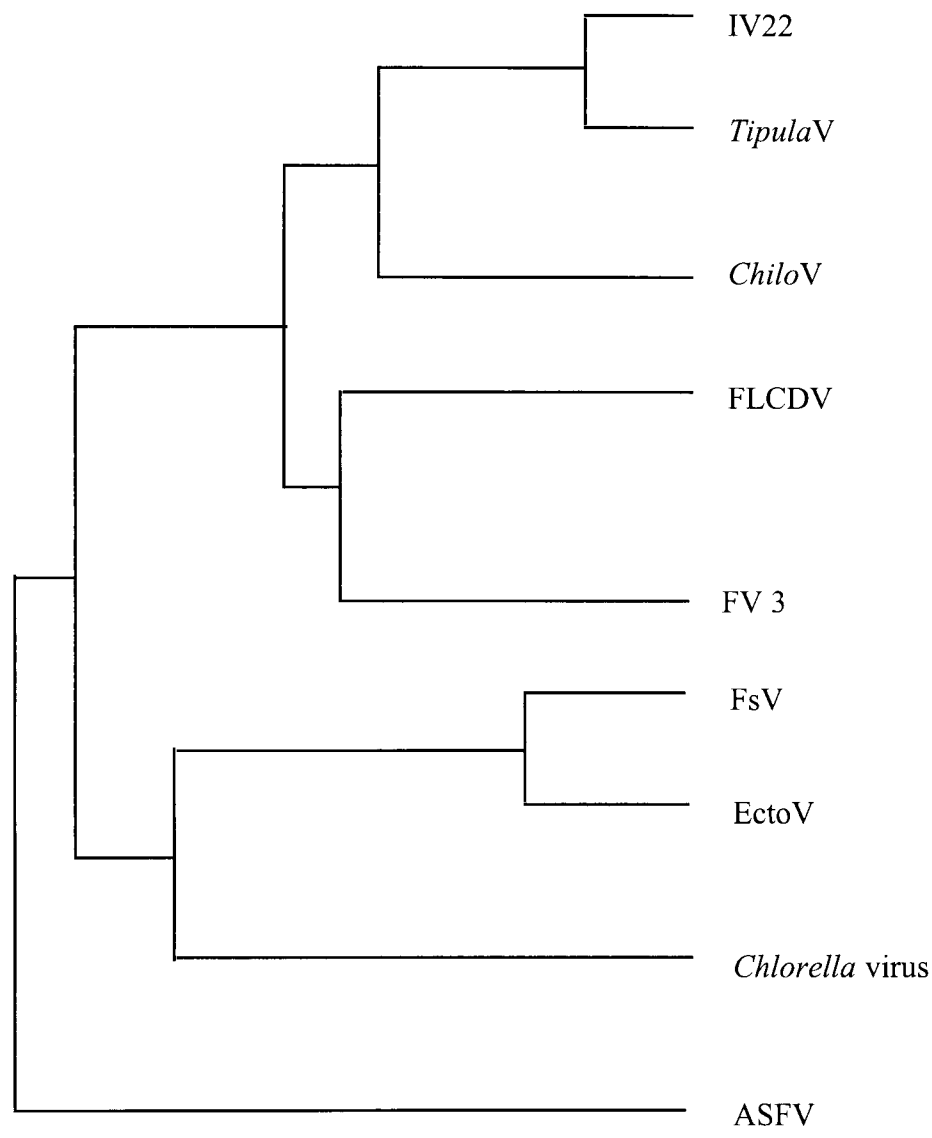

Fig. 8 The relationship among FsV, *Ectocarpus* virus-EctoV,*Chlorella* virus, iridoviruses (IV22, *Tipula* iridescent virus-TipulaV, *Chilo* iridescent virus-ChiloV), frog virus 3-FV3, fish lymphocystis disease viruse-FLCDV, and African swine fever virus-ASFV.

5. DISCUSSION and CONCLUSION

The primary goal of these experiments was to characterize the FsV transcripts from the persistently virus-infected sporophyte of *Feldmannia sp.* Because the infection is persistent in the alga, and so far virus-free *Feldmannia sp.* plants have not been found, it is not possible to determine the early or late transcripts by infecting the virus-free alga with virus particles. The strategy of using 4 cosmid clones that cover the entire viral genome to probe the Northern blots of *Feldmannia sp.* total RNA gives the general patterns of FsV transcripts: 1) there are 23 viral transcripts which were identified ranging from 0.9 kb to 5.5 kb; these transcripts probably only represent the most abundant transcripts or late transcripts; 2) viral transcriptional activity is not evenly distributed throughout the viral genome.

Two major transcripts that map to the 5.8 kb BamHI fragment -R were further characterized by DNA sequencing. According to the DNA sequence analysis no obvious eukaryotic promoters could be identified from the two genes. It is quite unusual that sequence upstream of the first AUG codon do not have an AT rich region nor a typical TATA box. These facts are unlike that of *Chlorella* virus PBCV-1, in which major capsid protein gene contains an AT-rich region upstream of the first AUG codon (Graves *et al.*, 1992, Schuster *et al.*, 1990). The TTTTTNT transcriptional termination motif found for vaccinia and PBCV-1 viruses is not seen in either gene. These unique characteristics of FsV suggest that FsV may employ some other mechanism in the regulation of initiation and termination of gene expression.

Although the 3′NTR (non-transcriptional region) of ORF1 has a common eukaryotic polyadenylation signal, AAUAAA, the Northern analysis and 3′RACE experiments showed that the 1.5kb major transcript is not polyadenylated. This is also true of the 0.9 kb transcript. In PBCV-1, the early transcripts usually are polyadenylated, whereas late transcripts usually are not polyadenylated (Schuster *et al.,* 1990). So these two genes may be late transcripts.

The ORF 1 gene product has significant similarity with the major capsid protein of *Chlorella* virus PBCV-1, the major capsid protein of Iridoviridae (*Tipula* iridescent virus, *Chilo* iridescent virus, IV22, frog virus 3, fish lymphocystis disease virus), and the major capsid protein of African swine fever virus that is closely related to the Iridoviridae family. In addition the ORF 1 contains three possible asparagine-linked glycosylation sites, which usually occur in large viral capsid proteins or structural proteins. So the ORF1 was identified as the gene for the major capsid protein (MCP) of FsV.

Considering the relatedness of FsV with other viruses, the major capsid protein of FsV has significant homology with the major capsid protein of *Chlorella* virus PBCV-1 (27% identity), this suggests that FsV is closely related to *Chlorella* virus. The major capsid protein of FsV also has homology with the major capsid protein of iridoviruses, FLCDV and ASFV in several conserved regions. The iridoviruses, FLCDV and ASFV, like FsV, are all icosahedral particles and contain a large double-stranded DNA genome (110-330 kb). These facts may suggest that the FsV is at least distantly related to these invertebrate and vertebrate viruses.

BIBLIOGRAPHY

Baker, J.R. and Evans, L.V. (1973) The ship fouling alga *Ectocarpus*. I ultrastructure and cytochemistry of plurilocular reproductive stages. Protoplasma 77:1-13

Cameron, I. (1990) Identification and characterization of the gene encoding the major structural protein of insect iridescent virus type 22. Virology 178:35-42

Chang, S. Puryear, and J. Cairney, J. (1993) A simple and efficient method for isolating RNA from pine trees. Plant Mol. Bio. Reporter. 11(2):113-116

Clitheroe, S.B. and Evans, L. V. (1974) Viruslike particles in the brown alga Ectocarpus. J. Ultrastr. Res. 49: 211-217

Flügel, R. M. (1985) Lymphocystis disease virus. Current Topics in Microbiology and Immunology. 116:133-150

Friess-Klebl, A. Knippers, R. and Müller, D. G. 1994. Isolation and characterization of a DNA virus infecting *Feldmannia simplex* (Phaeophyceae). J. Phycol. 30:653-658.

Graves, M. V. and Meints, R. H. (1992) Characterization of the major capsid protein and cloning of its gene from algal virus PBCV-1. Virology 188:198-207

Henry, E. C. and Meints, R. H. (1992) A persistent virus infection in *Feldmannia* (Phaeophyceae). J. Phycol. 28:517-526

Hu, M. C.-T. and Davidson, N. (1986) Mapping transcription start point on cloned genomic DNA with T4 DNA polymerase: a precise and convenient technique. Gene 42:21-29

Ivey, R.G. Henry, E.C. Lee, A.M. Klepper, L. Krueger, S.K. Meints, R.H.(1996) *Feldmannia* algal viruses have two genome-size classes: viral genome restriction site map. (in press)

Klein, M. Lanka, S.T.J. Knippers, R. and Müller, D. (1995) Coat protein of the Ectocarpus siliculosus virus. Virology 206:520-526

Krueger, S.K. Ivey, R.G. Henry, E.C. Meints, R.H. (1996) A brown algal virus genome contains a "RING" zinc finger motif. Viology 219:301-303

La Claire II, J. W. and West, J. A. (1977) Virus-like particles in the brown alga *Streblonema.* Protoplasma 93:127-130

Lanka, S. T. J., Klein, M., Ramsperger, U., Müller, D. G. and Knippers, R. (1993) Genome structure of a virus infecting the marine brown alga *Ectocarpus siliculosus.* Virology 193:802-811

Lee, M. A. Ivey, G. R. Henry, C. E. Meints, H. R. (1995) Characterization of a repetitive DNA element in a brown algal virus. Virology 211:474-480

Lopez-Otin, C. Freije, J. M. Parra, F. Mendez, E. and Vinuela, E. (1990) Mapping and sequence of the gene coding for protein p72, the major capsid protein of African Swine Fever virus. Virology 175:477-484

Markey, D.R. (1974) A possible virus infection in the brown alga *Phlaiella littoralis.* Protoplasma 80:223-232

McLachlan, J. (1973) Growth media-marine. In Stein, J. R. *Handbook of Phycological Methods. Culture Methods and Growth Measurements.* Cambridge University Press, New York, pp. 25-51

Müller, D. G., Kawai, H., Stache, B. and Lanka, S. (1990) A virus infection in the marine brown alga *Ectocarpus siliculosus* (Phaeophyceae). Botanica Acta 103: 72-82

Müller, D. G. and Stache, B. (1992) Worldwide occurrence of virus-infections in filamentous marine brown algae. Helgoländer Meeresunters 46:1-8

Müller, D. G. and Parodi, E. (1993) Transfer of a marine DNA virus from *Ectocarpus* to *Feldmannia* (Ectocarpales, Phaeophyceae): aberrant symptoms and restitution of the host. Protoplasma 175:121-125

Müller, D.G., Frenzer, K. (1993) Virus infections in three brown algae: *Feldmannia irregularis, F. simplex and Ectocarpus siliculosus*. Hydrobiologia 260:37-44

Oliveria , L. and Bisalputra, T. (1978) A virus infection in the brown alga *Sorocarpus uvarformis* (Lyngbye) *Pringsheim* (Phaeophyta, Ectocarpales) Ann. Bot. 42:439-445

Schnitzler, P. and Darai, G. (1993) Identification of the gene encoding the major capsid protein of fish lymphocystis disease virus. J. Gen. Virology 74:2143-2150

Schuster, A. M. Grave, M. Korth, K. Ziegelbein, M. Brumbaugh, J. Grone, D. Meints, R. (1990) Transcription and sequence studies of a 4.3-kb fragment from a ds-DNA eukaryotic algal virus. Virology 176:515-523

Skrdla, M.P.,Burbank, D.E., Xia, Y., Meints, R.H., and Van Etten, J.L. (1984) Structural proteins and lipids in a virus, PBCV-1, which replicates in a *Chlorella*-like alga. Virology 135:308-315

Stohwasser, R. Raab, K. Schnitzler, P. Janssen, W. and Darai, G. (1993) Identification of the gene encoding the major capsid protein of insect iridescent virus type 6 by polymerase chain reaction. J. Gen. Virology 74:873-879

Toth, R. and Wilce, R.T. (1972) Viruslike particles in the marine alga *Chorda tomentosa* (Laminariales). J. Phycol. 10:170-185

Van Etten, J. L. Lane, L. C. Meints, R. H. (1991) Viruses and viruslike particles of eukaryotic algae. Microbiol. Rev. 55:586-620